On Computing the Resultant of Generic Bivariate Polynomials

Gilles Villard

On Computing the Resultant of Generic Bivariate Polynomials

GILLES VILLARD, Univ Lyon, CNRS, ENS de Lyon, Inria, Université Claude Bernard Lyon 1, LIP UMR 5668, F-69007 Lyon, France

An algorithm is presented for computing the resultant of two generic bivariate polynomials over a field K. For such p and q in $\mathsf{K}[x, y]$ both of degree d in x and n in y, the algorithm computes the resultant with respect to y using $(n^{2-1/\omega}d)^{1+o(1)}$ arithmetic operations in K, where two $n \times n$ matrices are multiplied using $O(n^\omega)$ operations. Previous algorithms required time $(n^2 d)^{1+o(1)}$.

The resultant is the determinant of the Sylvester matrix $S(x)$ of p and q, which is an $n \times n$ Toeplitz-like polynomial matrix of degree d. We use a blocking technique and exploit the structure of $S(x)$ for reducing the determinant computation to the computation of a matrix fraction description $R(x)Q(x)^{-1}$ of an $m \times m$ submatrix of the inverse $S(x)^{-1}$, where $m \ll n$. We rely on fast algorithms for handling dense polynomial matrices: the fraction description is obtained from an x-adic expansion via matrix fraction reconstruction, and the resultant as the determinant of the denominator matrix.

We also describe some extensions of the approach to the computation of generic Gröbner bases and of characteristic polynomials of generic structured matrices and in univariate quotient algebras.

1 INTRODUCTION

Given two bivariate polynomials p and q in $\mathsf{K}[x, y]$ where K is a commutative field, we address the question of computing the resultant $\mathrm{Res}_y(p, q) \in \mathsf{K}[x]$ of p and q with respect to y. We take $p = \sum_{i=0}^{n} p_i(x)y^i$ and $q = \sum_{i=0}^{n} q_i(x)y^i$ of degree d in x and n in y. The Sylvester matrix $S(x) \in \mathsf{K}[x]^{2n\times 2n}$ associated to p and q is defined by $s_{i,j} = p_{n+j-i}$ and $s_{i,j+n} = q_{n+j-i}$, for $1 \le i \le 2n$ and $1 \le j \le n$ (where $p_k = q_k = 0$ when $k < 0$ or $k > n$). The resultant of p and q with respect to y is the determinant of S.

Since the early 1970's it is known that the bivariate resultant can be computed in $(n^2 d)^{1+o(1)}$ arithmetic operations in K. This complexity bound is obtained by combining an evaluation-interpolation approach à la Collins and Brown for the multivariate resultant [8, 10], and the half-gcd algorithm resulting from the works of Knuth [30] and Schönhage [45] for the integers, and of Moenck [39] for the univariate polynomials. We refer to [16, Chap. 11] for more details and references. More precisely, on the one hand, the resultant of two univariate polynomials of degree n (taking $d = 0$ in above definition) can be computed in $O(\mathsf{M}(n)\log n)$ arithmetic operations in K using the Knuth-Schönhage-Moenck algorithm. We use $\mathsf{M}(n)$ for a multiplication time for univariate polynomials of degree bounded by n over K (see for instance [16, Chap. 8]). On the other hand, in our case the resultant has degree at most $2nd$, hence an extra factor nd appears for the evaluation-interpolation cost. In total, it can be shown that the bivariate resultant can be computed using $O(n\,\mathsf{M}(nd)\log(nd))$ arithmetic operations [16, Chap. 11], which is $(n^2 d)^{1+o(1)}$ using $\mathsf{M}(n) = O(n \log n \log\log n)$ with Cantor and Kaltofen's polynomial multiplication [9].

Before giving an overview of our approach let us mention some important results that have been obtained since the initial results cited above. For comprehensive presentations of the resultant and subresultant problem, and detailed history and complexity analyses, the reader may refer to [16, 17, 36]. Especially for avoiding modular methods over $\mathbb{Z}$,

Author's address: Gilles Villard, Univ Lyon, CNRS, ENS de Lyon, Inria, Université Claude Bernard Lyon 1, LIP UMR 5668, F-69007 Lyon, France.

recursive subresultant formulas have been given in [17, 38, 43] that allow half-gcd schemes for computing the resultant of polynomials in $\mathsf{D}[y]$ where D is a domain such that the exact division can be performed.

The complexity bound $(n^2d)^{1+o(1)}$ has not been improved in the general case. In some special cases much better complexity bounds are known [5, Sec. 5]. In particular, for univariate f and g of degree n in $\mathsf{K}[y]$, the composed sum $(f \oplus g)(x) = \mathrm{Res}_y(f(x-y), g(y))$ and the composed product $(f \otimes g)(x) = \mathrm{Res}_y(y^n f(x/y), g(y))$ can be computed using $n^{2+o(1)}$ operations in K [5]. (The restrictions in [5] for fields of small characteristic may be bypassed by techniques such as those in [33] and references therein.) The latter bound improves upon $n^{3+o(1)}$ (taking $d = n$, $p(x, y) = f(x - y)$, and $q(x, y) = g(y)$) and is essentially optimal since the composed sum and product have degree n^2.

Another special bivariate case concerns particular resultants that are linear with respect to one of the variables. The question is related to characteristic polynomials of special structured matrices. For a and g monic of degree n in $\mathsf{K}[y]$, the characteristic polynomial of a in the quotient algebra $\mathcal{A} = \mathsf{K}[y]/\langle g(y)\rangle$ is given by $\chi(x) = \mathrm{Res}_y(x - a(y), g(y))$ (see for instance [13, Chap. 4, Proposition (2.7)]). Representing the endomorphism of the multiplication by a by an appropriate $n \times n$ matrix A, χ is the characteristic polynomial $\det(xI - A)$ of A. Let ω be the exponent of fast matrix multiplication. In [46, Theo. 3.4] (see also [47] and [28, Sec. 6]), with a formulation that uses mainly polynomials, it is shown that an algorithm exists for computing the minimal polynomial μ of a using $n^{1.5+o(1)} + O(n^{(\omega+1)/2})$ operations in K. Using $\omega = 2.373$ [12, 34], the latter bound is $O(n^{1.687})$. The resultant point of view is adopted in [5] where an approach using similar tools than in [46, 47] leads to the same exponent for the characteristic polynomial, hence the special resultant χ (see above for fields of small characteristic). With $\omega < 3$, the bound $O(n^{(\omega+1)/2})$ is better than $n^{2+o(1)}$ for a general resultant (taking $d = 1$, $p(x, y) = x - a(y)$, and $q(x, y) = g(y)$).

Both methods in [46] and [5] use a K-linear map $\pi : \mathcal{A} \to \mathsf{K}$. The map in [46] allows to reduce the minimal polynomial problem to the one of computing generators of linearly generated sequences [44, 50]. The trace is used in [5], and allows via Le Verriers's approach to compute the characteristic polynomial using Newton identities. With a linear algebra point of view we note that those latter algorithms could compare to a structured version of the one in [26]. Using Wiedemann's method [53], the minimal polynomial computation can be reduced to the computation of generators of linearly generated sequences with $\pi : A \mapsto X^T AY \in \mathsf{K}$ where X and Y are vectors.

We also mention the particular situation where only a few terms of the resultant are needed. The algorithm of [40] computes the truncated resultant $\mathrm{Res}_y(p, q) \bmod x^k$ in $(nk)^{1+o(1)}$ operations in K. The latter bound improves upon a division-free computation of the resultant over $\mathsf{K}[x]/\langle x^k\rangle$ in time $(n^2k)^{1+o(1)}$. For a division-free univariate resultant over K in $n^{2+o(1)}$ operations one can indeed apply Strassen's removal of divisions to polynomials defined by: $p_0(x) = 1$, $p_1(x) = x$, and $p_k(x) = xp_{k-1}(x) + p_{k-2}(x)$ for $k \geq 2$[1].

Our contribution. The complexity bound $(n^2d)^{1+o(1)}$ for the bivariate resultant is roughly speaking the product of the essentially linear time for the half-gcd over K by the degree $O(nd)$ of the answer. For generic input polynomials p and q, we reduce the complexity bound below this product. We are going to prove the following theorem.

Theorem 1.1. (Proven in Section 6.) *Let p and q in $\mathsf{K}[x, y]$ be of degree d in x and n in y. Except if (p, q) is on a certain hypersurface of $\mathsf{K}^{2(n+1)(d+1)}$ the resultant $\mathrm{Res}_y(p, q) \in \mathsf{K}[x]$ of p and q can be computed using $(n^{2-1/\omega}d)^{1+o(1)}$ operations in K*[2,3].

[1] The author thanks E. L. Kaltofen for showing this construction.
[2] One should expect a slight exponent improvement by employing fast rectangular matrix multiplication [11, 22, 35].
[3] Links to Maple worksheets with some constructions of the paper and examples are provided on the page http://perso.ens-lyon.fr/gilles.villard/mws/issac18.

For two generic polynomials p and q of degree n the best known bound was $n^{3+o(1)}$, we obtain $n^{8/3+o(1)}$ with $\omega = 3$, and $O(n^{2.58})$ asymptotically. Our approach can be used for other types of structured matrices, such as those with small displacement rank [25]. For example, an immediate consequence is the computation of the characteristic polynomial of a Toeplitz matrix (see Section 7). The best known bound was $n^{2+o(1)}$ [42]. We obtain $n^{5/3+o(1)}$ or $O(n^{1.58})$ for the characteristic polynomial of a generic Toeplitz matrix. Another case linear in x has been mentioned above: characteristic polynomials in univariate quotient algebras can also be computed in a generic case in $O(n^{1.58})$ operations, which improves upon previous methods whose cost is $O(n^{1.687})$. Note that, already for $\omega = 3$, the bound $n^{5/3+o(1)}$ improves upon $n^{2+o(1)}$. In view of the above discussion concerning the algorithms in [5, 26, 46] we are going to see that our approach may be seen as using several maps simultaneously, somewhat in the spirit of the blocked version [29] of [26].

Overview of the approach. An often successful idea for computing the determinant of a matrix M over $\mathsf{K}[x]$ or $\mathbb{Z}$, is to reduce the problem using Cramer's rule to the solution of one or a few linear systems $M^{-1}y$ for well chosen vectors y [41]. For example in the integer case, the determinant (or, for a non generic M, the largest invariant factor) can be recovered from the denominators of the entries of a few random system solutions [1, 15, 49]. The latter approach works in three phases: (i) use lifting for computing truncated p-adic expansions of system solutions $M^{-1}y$; (ii) reconstruct corresponding (scalar) fractions for the entries of the solutions; (iii) deduce the determinant (or the largest invariant factor) of M from denominators. The polynomial matrix case is studied in [48] and follows analogous steps starting instead with x-adic expansions.

Our resultant algorithm is given in Section 6. It generalizes the three phases (i)-(ii)-(iii) for computing the determinant of the Sylvester matrix $S(x)$, and some major changes are required. The improvement is obtained thanks to: a block approach—instead of linear system solving; the expansion and reconstruction of matrix polynomial fractions—instead of scalar polynomial fractions; the replacement of system solution lifting by a specific expansion phase that takes advantage of the structure of the input matrix.

Indeed, reducing the determinant problem to system solution is an appropriate strategy in many cases, however, given a $2n \times 2n$ Sylvester matrix $S(x)$ of degree d, a linear system solution $S(x)^{-1}y$ will have size $\Omega(n^2 d)$ in a generic sense, which is too large for the objective of improving the resultant complexity bound. Inspiration may be gained from block Krylov subspace techniques—see [29] and references therein—for the determinant of a characteristic matrix $M = xI - A$. Rather than fully solving systems $S(x)^{-1}y$, a first idea is to circumvent the difficulty by computing only several entries of several linear system solutions. "Several" here means $m \ll n$ where m is chosen as a function of n in Section 6 for minimizing the overall cost. More precisely, we consider the $m \times m$ north-eastern submatrix

$$H(x) = X^T S(x)^{-1} Y \in \mathsf{K}(x)^{m \times m}$$

of $S(x)^{-1}$, where $X = [I_m, 0, \ldots, 0]^T$ and $Y = [0, 0, \ldots, I_m]^T$ are $(2n) \times m$ matrices, and I_m is the identity matrix of dimension m. We are going to see that generically this $m \times m$ submatrix of the inverse of $S(x)$ has a (right) fraction description (among other properties, see further below)

$$H(x) = R(x)Q(x)^{-1} \tag{1}$$

that can be computed fast, where R and Q are polynomial matrices of dimension m and degree at most $2\lceil n/m \rceil d$. The necessary background on fraction descriptions will be given in Section 2. Our choice of X and Y will in particular simplify the presentation for Gröbner bases in Section 7. Different and more general choices could however be made. For instance, with $X = Y = [0, 0, \ldots, I_m]^T$ the fraction H could be seen as the inverse of a Schur complement.

Once the description (1) is available the resultant is deduced from the denominator matrix (generalization of Cramer's rule, see Lemma 2.1) since generically we will have

$$\mathrm{Res}_y(p,q)(x) = \det S(x) = s \det Q(x), \text{ for some } s \in \mathsf{K} \setminus \{0\}.$$

The leading term of the resultant, depending here on the scalar s, will be computed separately. In order to prove the degree bound on R and Q, we first identify in Section 3 special polynomials $\bar{p}$ and $\bar{q}$ whose Sylvester matrix leads to a (sufficiently) generic degree behaviour. We then show in Section 4 that the generic behaviour corresponds to (p,q)'s not on a certain hypersurface. The polynomials $\bar{p}$ and $\bar{q}$ will be $\bar{p}(x,y) = x^d y^n + y^m$ and $\bar{q}(x,y) = y^n + x^d$, where $1 \le m \le n$. For example, with input degrees $n = 8$ and $d = 1$, and blocking factor $m = 3$, consider the polynomials $\bar{p}(x,y) = xy^8 + y^3$ and $\bar{q}(x,y) = y^8 + x$. The 3×3 north-eastern submatrix of the inverse of their 16×16 Sylvester matrix $\bar{S}(x)$ satisfies

$$\bar{H}(x) = \bar{R}(x)\bar{Q}(x)^{-1} = -\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x & 0 & x^4 \\ x^6 & x & 0 \\ 0 & x^6 & x \end{bmatrix}^{-1},$$

with $\bar{Q}$ of degree $2\lceil 8/3 \rceil = 6$, and we have $\mathrm{Res}_y(xy^8 + y^3, y^8 + x) = \det \bar{Q}(x) = x^{16} + x^3$.

The size of the description RQ^{-1} will be generically $O(mnd)$, hence smaller than the size of a system solution, hence the description can be handled with fewer operations. Furthermore, in the same Section 4 we also make use of a small (left) description $H = Q_L^{-1} R_L$. The existence of both a small left and a small right description implies that only $O(nd/m)$ terms of an x-adic expansion of H suffice for computing the denominator Q using matrix fraction reconstruction (see Lemma 2.5 and Corollary 4.2).

Another issue then arises for computing the initial x-adic expansion of H. The immediate approach that would compute an expansion of $S(x)^{-1}Y$ and in the end keep only $XS(x)^{-1}Y$ would have prohibitive cost. Computing only selected entries of the inverse therefore requires a modification of the expansion step. We circumvent the difficulty by making use of the Toeplitz-like structure of the Sylvester matrix. One will find the necessary reminders concerning dense structured matrices in Section 5. The reader may refer for example to [3, 21] for an insight into the subject. As well as S, the inverse S^{-1} is Toeplitz-like. With our choice of X and Y what follows is essentially the fact that H, as a special submatrix of S^{-1}, is also Toeplitz-like. One can write S^{-1} in ΣLU form [25, 27], we mean:

$$S(x)^{-1} = \sum_{i=1}^{2} L_i(x) U_i(x) \in \mathsf{K}(x)^{(2n)\times(2n)} \tag{2}$$

for some lower triangular Toeplitz matrices L_1 and L_2 and some upper triangular Toeplitz matrices U_1 and U_2. We will then use the fact that thanks to the triangular structure of L_1 and L_2 one has

$$H(x) = \sum_{i=1}^{2} L_i^{(m)}(x) T_i^{(m)}(x) \in \mathsf{K}(x)^{m\times m} \tag{3}$$

where the $L_i^{(m)}$'s and $T_i^{(m)}$'s are Toeplitz submatrices of the L_i's and U_i's. A key point is that from the ΣLU representation of S^{-1} the computation of H will involve two multiplications of Toeplitz matrices of dimension only $m \times m$.

Structured matrix inversion has been well studied. Over a field, a randomized divide-and-conquer approach allows to compute—via Schur complements—a ΣLU representation of the inverse of a Toeplitz-like matrix using $O(\mathsf{M}(n)\log n)$ arithmetic operations, see [4, 27] and [3, Chap. 2, Sec. 13]. The Sylvester matrix is block Toeplitz (after a colum permutation, with 1×2 rectangular blocks), which is a special Toeplitz-like case. Therefore we will rather follow the path of [31] for reducing the problem to matrix Padé approximation and to the use of the half-gcd algorithm.

For the sake of completeness we propose a specific study of the Sylvester matrix case in Section 5. We use known techniques for deriving an explicit formula for a ΣLU representation (2) of its inverse over a field. Such a formula allows to compute a ΣLU representation essentially via the half-gcd algorithm in time $O(\mathsf{M}(n)\log n)$. With operations on truncated power series a corresponding expansion of order $O(nd/m)$ is therefore obtained using $(n^2d/m)^{1+o(1)}$ operations. Finally, taking advantage of Toeplitz structures in (3), an appropriate expansion of the $m \times m$ matrix H is obtained via Toeplitz matrix multiplication using $(m^2 \times nd/m)^{1+o(1)}$ operations. For instance, taking $m = \sqrt{n}$, we see that computing an order $O(\sqrt{n}d)$ expansion of a $\sqrt{n} \times \sqrt{n}$ submatrix of $S(x)^{-1}$ (hence n entries) has cost $(n^{3/2}d)^{1+o(1)}$, while solving a linear system (n entries also) via an expansion of order $O(nd)$ would require $(n^2d)^{1+o(1)}$ operations.

Dense linear algebra for reconstructing $H = RQ^{-1}$ from its expansion [2, 19] and for computing the determinant of Q [32, 48] can be performed using $(m^\omega \times nd/m)^{1+o(1)}$ operations. We will see in Section 6 that it follows that the overall cost for the resultant is minimized by choosing $m \sim n^{1/\omega}$.

Some consequences: generic bivariate ideals and characteristic polynomials. In Section 7 we discuss some corollaries and extensions of the above approach. For instance, from the structure of the Sylvester matrix, we notice that the entries of Q give the coefficients in $\mathsf{K}[x]$ of polynomials in the ideal $\mathcal{I}$ generated by p and q. Those polynomials have degree less than m in y, and a Gröbner basis of $\mathcal{I}$ can be deduced for the lexicographic order within the complexity bound of Theorem 1.1. We will also discuss the case of polynomials linear in one of the variables and compute characteristic polynomials.

2 MATRIX FRACTIONS

We give the basic notions and results we need in the rest of the text concerning matrix fraction descriptions. The reader may refer to the comprehensive material in [24]. By analogy with scalar polynomial fractions, an $n \times m$ rational matrix $H(x)$ over $\mathsf{K}(x)$ can be written as a fraction of two polynomial matrices. A right fraction description is given by $R(x) \in \mathsf{K}[x]^{n\times m}$ and $Q(x) \in \mathsf{K}[x]^{m\times m}$ such that

$$H(x) = R(x)Q(x)^{-1} \in \mathsf{K}(x)^{n\times m},$$

and a left description by $R_L(x) \in \mathsf{K}[x]^{n\times m}$ and $Q_L(x) \in \mathsf{K}[x]^{n\times n}$ such that

$$H(x) = Q_L(x)^{-1}R_L(x) \in \mathsf{K}(x)^{n\times m}.$$

Degrees of denominator matrices are minimized using column reduced forms. A non singular polynomial matrix is said to be column reduced if its leading column coefficient matrix is non singular [24, Sec. 6.3].

This leads us to the definition of *irreducible and minimal fraction descriptions.* If R and Q (resp. R_L and Q_L) have unimodular right (resp. left) matrix gcd's [24, Sec. 6.3] then the description is called irreducible. If in addition Q (resp. Q_L) is column reduced then the description is called minimal. Our determinant algorithm relies on the following lemma that we will use as a generalization of Cramer's rule.

Lemma 2.1. ([24, Lemma 6.5-9].) *The denominators of irreducible matrix fraction descriptions all have the same non unity invariant factors (non unity entries in the Smith normal form). They all have in particular the same determinant up to a non zero element of* K.

We also have a multiplicative property for the denominators.

Lemma 2.2. ([24, Lemma 6.5-5].) *Let $Q \in \mathsf{K}[x]^{m\times m}$ be the denominator of an irreducible right description of H and $Q' \in \mathsf{K}[x]^{m\times m}$ be the denominator of an arbitrary right description of H. There exists a polynomial matrix M such that $Q' = QM$.*

For a non singular square polynomial matrix M, we now study the special fraction $H = X^T M(x)^{-1} Y$, where for $1 \le m \le n$, $X = [I_m, 0, \ldots, 0]^T$ and $Y = [0, 0, \ldots, I_m]^T$ are in $\mathsf{K}^{n\times m}$. Let M_{NE} be the square submatrix of dimension $2n - m$ given by rows $1, \ldots, 2n - m$ and columns $m + 1, \ldots, 2n$ of M, let also M^* be the adjoint matrix of M. From the Schur complement formula, $g(x) = \det M_{\text{NE}}(x) = (\det(X^T M^*(x) Y))/(\det M(x))^{m-1}$ is a polynomial in $\mathsf{K}[x]$. Then if $X^T M^{-1} Y = RQ^{-1}$, taking determinants in both sides we obtain

$$g(x)/\det M(x) = \det R(x)/\det Q(x). \tag{4}$$

Hence $\det Q$ divides $\det M$ whenever RQ^{-1} is irreducible and we have the next lemma.

Lemma 2.3. (Compare with [29, Theorem 2.12].) *The denominators of irreducible matrix fraction descriptions of $X^T M^{-1} Y$ all have same determinant up to a non zero element of K, which is a divisor of $\det M$.*

The next thing we need is a characterization of the link between the degree of fraction descriptions and the number of terms required for reconstructing the fraction from its expansion. In the matrix case the number of terms depends on both the left and right degrees since those degrees may be different. A matrix fraction is said to be strictly proper if it tends to zero when x tends to infinity. The material of next lemma can be found along the lines of [24], or [52, Sec. 4] and [29, Sec. 2]. The two latter references illustrate the links of our algorithm with block Krylov methods.

Lemma 2.4. *Let $H \in \mathsf{K}(x)^{m\times m}$ be a strictly proper fraction, and write $H(x) = \sum_{k\ge 0} H_k x^{-1-k}$. For any integer $\delta \ge 0$, let also $\mathcal{H}$ be the following block Hankel matrix:*

$$\mathcal{H} = \begin{bmatrix} H_0 & H_1 & \cdots & H_{\delta-1} \\ H_1 & \iddots & \iddots & H_\delta \\ \vdots & \iddots & \iddots & \vdots \\ H_{\delta-1} & H_\delta & & H_{2\delta-2} \end{bmatrix} \in \mathsf{K}^{(m\delta)\times(m\delta)},$$

and denote by $\mathcal{H}_\infty$ the corresponding infinite block Hankel matrix. The rank of $\mathcal{H}_\infty$ is the determinantal degree ν of denominators of irreducible fraction descriptions of H; $\operatorname{rank}\mathcal{H} = \operatorname{rank}\mathcal{H}_\infty$ if and only if H has left and right fraction descriptions of degree at most δ.

Proof. *(Brief outline.)* One has a correspondence between denominators of descriptions and minimum generators of the sequence $\{H_i\}_{i\ge 0}$ [29, Lemma 2.8]. The fact that the rank of $\mathcal{H}_\infty$ is ν is by applying [29, (2.6)] both on the left and right side, as well as the fact that $\operatorname{rank}\mathcal{H} = \nu$ as soon as descriptions have degree at most δ. Conversely, if $\operatorname{rank}\mathcal{H} = \operatorname{rank}\mathcal{H}_\infty$ then, arguing in a similar way than in [29, Cor. 3.8], relations between columns (resp. rows) in $\mathcal{H}_\infty$ provide right (resp. left) descriptions of degree at most δ. □

We will use Lemma 2.4 for showing in the next two sections that small degree descriptions exist in the generic case. Among the methods for matrix fraction reconstruction from the expansion of H (see for instance [29, Sec. 2]), the proof of Lemma 2.4 indicates that one could use block-Hankel system solving. We will rather use a deterministic approach based on polynomial matrices and order bases computation [2, 19].

Generically we will be in the situation where $H \in \mathsf{K}(x)^{m\times m}$ has a power series expansion. An order basis (also known as minimal approximant basis or σ-basis) [2] of $[H(x) \ -I_m] \in \mathsf{K}[[x]]^{m\times(2m)}$ with order σ, is a minimal (column reduced) basis of the module of vectors $u \in \mathsf{K}[x]^{2m}$ such that $[H(x) \ -I_m]\, u(x) \equiv 0 \mod x^\sigma$. In particular, an order basis is a polynomial matrix in $\mathsf{K}[x]^{(2m)\times(2m)}$.

Lemma 2.5. Matrix fraction reconstruction. *([19, Lemma 3.7].) Let $H \in \mathsf{K}(x)^{m\times m}$ be a strictly proper power series, with left and right matrix fractions descriptions of degree at most δ. The m columns $[Q^T \ R^T]^T$ of degree at most δ of an order basis of $[H \ -I_m]$ with order $2\delta+1$ define a minimal description RQ^{-1} of H.*

We use order bases in a special case. For more recent results on the problem the reader may refer to [23, 54].

3 A SPECIAL SYLVESTER MATRIX

For polynomials p and q in $\mathsf{K}[x,y]$ with non singular Sylvester matrix S, we consider the $m\times m$ north-eastern submatrix H of S^{-1}. In this section we construct special polynomials $\bar{p}$ and $\bar{q}$ such that the corresponding $\bar{H}$ has small degree—we mean in $O(nd/m)$—minimal descriptions; the latter having denominators whose determinant is equal to the one of the Sylvester matrix $\bar{S}$ of $\bar{p}$ and $\bar{q}$ up to a non zero constant. Note that we know by Lemma 2.3 that the determinant of such a denominator is a divisor of the resultant. The properties of $\bar{p}$ and $\bar{q}$ will allow us in the next section to identify the degree behaviour for descriptions associated to generic polynomials.

For integers $l, m \geq 1$, consider $B \in \mathsf{K}[x]^{l\times(l+m)}$ such that

$$b_{i,i} = x^d \text{ and } b_{i,i+m} = -1 \text{ for } 1 \leq i \leq l, \tag{5}$$

and $b_{i,j} = 0$ otherwise. We construct m vectors that form the columns of a nullspace basis $P \in \mathsf{K}[x]^{(l+m)\times m}$ of B. Among those vectors, m_1 are of degree $\delta = \lceil l/m\rceil d$ and $m_2 = m\lceil l/m\rceil - l$ are of degree $\delta - d = \lfloor l/m\rfloor d$, such that the sum of the degrees is ld. If m divides l then the m vectors are of degree $\delta = ld/m$. Those vectors have zero entries but for $p_{k,k} = 1, 1 \leq k \leq m$, and in each column $1 \leq j \leq m$ for $p_{k+m,j} = x^d p_{k,j}, 1 \leq k \leq l$. The columns of P are in the nullpace of B, one can check that $\sum_k b_{i,k}p_{k,j} = 0$. Indeed if for some ξ, $b_{i,\xi} = x^d$ then the only other non zero entry in row i is $b_{i,\xi+m} = -1$, and $\sum_k b_{i,k}p_{k,j} = b_{i,\xi}p_{\xi,j} + b_{i,\xi+m}p_{\xi+m,j} = x^d p_{\xi,j} - p_{\xi+m,j} = 0$. Then we note that P is a basis of the nullspace since it admits I_m as uppermost submatrix, hence it cannot be divided by a non unimodular matrix. The entries of P with highest column degrees are in its lowest submatrix, which is say $D_{\delta,d}(x) = \mathrm{diag}(x^\delta, \ldots, x^\delta, x^{\delta-d}, \ldots, x^{\delta-d}) \cdot C$ for a C column permutation:

$$P(x) = \begin{bmatrix} I_m \\ \text{degrees} < \delta \text{ or } \delta - d \\ D_{\delta,d}(x) \end{bmatrix} \in \mathsf{K}[x]^{(l+m)\times m}. \tag{6}$$

We then consider $\bar{p}(x,y) = x^d y^n + y^m$ and $\bar{q}(x,y) = y^n + x^d$, $1 \leq m \leq n$. The corresponding Sylvester matrix is

$$\bar{S}(x) = \begin{bmatrix} x^d I_n + I'_{n|m-n} & I_n \\ I'_{n|m} & x^d I_n \end{bmatrix} \in \mathsf{K}[x]^{(2n)\times(2n)}, \tag{7}$$

where $I'_{n|k} \in \mathsf{K}^{n\times n}$ has 1's on the upper diagonal starting at entry $(1, k+1)$ if $k \geq 0$, or on the lower diagonal starting at entry $(-k+1, 1)$ otherwise. We then introduce P_1 such that $\bar{S}P_1 \in \mathsf{K}[x]^{(2n)\times n}$ is:

$$\bar{S}(x) \begin{bmatrix} -I_n \\ x^d I_n + I'_{n|m-n} \end{bmatrix} = \begin{bmatrix} 0 \\ x^{2d} I_n - I'_{n|m} + x^d I'_{n|m-n} \end{bmatrix}.$$

Rows $n+1$ to $2n-m$ of $\bar{S}P_1$ form a matrix as in (5) with $l = n-m$, whose nullspace is described by a matrix P_2 as in (6), with $\delta_0 = 2\lceil (n-m)/m \rceil d$. The first $2n-m$ rows of $\bar{S}$ can therefore be annihilated using:

$$P_1(x)P_2(x) = \begin{bmatrix} -I_m \\ \text{degrees} < \delta_0 + d \text{ or } \delta_0 - d \\ D_{\delta_0+d,2d}(x) + I_m \end{bmatrix} \in \mathsf{K}[x]^{(2n)\times m}.$$

Now, applying P_1P_2 to $\bar{S}$ given by (7) we arrive at:

$$\bar{S}(x)P_1(x)P_2(x) = \begin{bmatrix} 0 \\ D_{\delta_0+2d,2d}(x) + I_m \end{bmatrix} \in \mathsf{K}[x]^{(2n)\times m}.$$

Taking $\delta = \delta_0 + 2d$ and $\bar{Q}(x) = D_{\delta,2d}(x) + x^d I_m \in \mathsf{K}[x]^{m\times m}$, with

$$X = \begin{bmatrix} I_m, 0, \ldots, 0 \end{bmatrix}^T, \; Y = \begin{bmatrix} 0, \ldots, 0, I_m \end{bmatrix}^T \in \mathsf{K}^{(2n)\times m}, \tag{8}$$

we are led to $P_1(x)P_2(x) = \bar{S}(x)^{-1}Y\bar{Q}(x)$. This gives

$$\bar{H}(x) = X^T\bar{S}(x)^{-1}Y = -\bar{Q}(x)^{-1} \in \mathsf{K}(x)^{m\times m}. \tag{9}$$

The right and left descriptions $\bar{Q}^{-1} = -I_m\bar{Q}^{-1} = -\bar{Q}^{-1}I_m$ are irreducible, therefore by Lemma 2.3 $\det\bar{Q}$ is a divisor of $\det\bar{S}$. Both determinants have leading coefficient ± 1, and $\deg\det\bar{Q} = 2nd$. In terms of determinants (4) and (9) can be rewritten as

$$\det(X^T\bar{S}(x)^{-1}Y) = \pm 1/\det\bar{Q}(x) = \pm 1/\det\bar{S}(x). \tag{10}$$

As announced at the beginning of the section, $\bar{Q}$ has degree in $O(nd/m)$ and its determinant gives the resultant.

Writing $\bar{H}(x) = \sum_{k\geq 0}\bar{H}_k x^{-1-k}$, we consider the Hankel matrices $\bar{\mathcal{H}} \in \mathsf{K}^{(m\delta)\times(m\delta)}$, where $\delta = 2\lceil n/m \rceil d$, and $\bar{\mathcal{H}}_\infty$ as in Lemma 2.4. Applying the latter lemma, since we have left and right descriptions of degree at most δ we know that

$$\operatorname{rank}\bar{\mathcal{H}} = \deg\det\bar{Q} = 2nd. \tag{11}$$

We have proceeded by block elimination for studying submatrices of S^{-1}. Direct inversion could have been used also.

4 DEGREES IN THE GENERIC CASE

Using the special Sylvester matrix of previous section, we now show that submatrices of S^{-1} have small descriptions generically. We consider generic bivariate polynomials p and q of degree d in x and n in y, we mean whose coefficients are distinct indeterminates $\alpha_{i,j}$ and $\beta_{i,j}$, for $0 \leq i \leq n$ and $0 \leq j \leq d$. We denote the polynomial ring $\mathsf{K}[\alpha_{0,0}, \ldots, \alpha_{n,d}, \beta_{0,0}, \ldots, \beta_{n,d}]$ by $\mathcal{R}_{\alpha,\beta}$, and the corresponding field of fractions by $\mathcal{F}_{\alpha,\beta}$. By genericity, the Sylvester matrix S associated to p and q is non singular (see also Φ_2 further below). Let Q be the denominator of a minimal description such that:

$$H(x) = X^TS(x)^{-1}Y = R(x)Q(x)^{-1}, \tag{12}$$

where X and Y are as in (8).

We first show that $\det Q$ is $\det S$ up to the leading coefficient. Then we show that Q has small column degrees, we mean in $O(n/m)$, hence Q can be computed from $O(n/m)$ terms of an expansion of H by Lemma 2.5. We may draw a parallel with the study of "lucky" projections X and Y in Coppersmith's block Wiedemann algorithm [29, 51]: unlucky projections may cause a drop in the determinantal degree of the minimal sequence generator, and/or increase the required length of the sequence.

Using Lemma 2.3 we know that $\det Q$ is a divisor of $\det S$. The fractions in (4) with $M = S$ are irreducible since they are irreducible in the special case (10). It follows that $\det Q$ is $\det S$ up to the leading term, and the polynomial $\Phi_1 = \mathrm{Res}_x(g, \det S) \in \mathcal{R}_{\alpha,\beta}$, where g has been define at (4), is non identically zero. Next, write $H(x) = \sum_{k\geq 0} H_k x^{-1-k}$, and, with $\delta = 2\lceil n/m\rceil d$ used for (11), consider the Hankel matrix $\mathcal{H} \in \mathcal{F}_{\alpha,\beta}^{(m\delta)\times(m\delta)}$ as in Lemma 2.4. Defining $\Phi_2 \in \mathcal{R}_{\alpha,\beta}$ to be the determinant of the coefficient matrix of degree d of S, the entries of $\mathcal{H}$ can be written as fractions with denominators being powers of Φ_2. In particular, $\mathcal{H}$ is well defined for $\bar{S}$ in (7) whose coefficient of degree d is the identity. Therefore Φ_2 is non trivial, which implies that H is strictly proper and $\deg\det S = 2nd$. Finally using (11) we know that $\mathrm{rank}\,\mathcal{H} \geq 2nd$, and $\mathrm{rank}\,\mathcal{H} = 2nd$ since $2nd$ is the maximum possible value. Therefore, $\mathrm{rank}\,\mathcal{H}_\infty = \mathrm{rank}\,\mathcal{H}$, and by Lemma 2.4 the fraction H has descriptions of degree at most δ . Let $\Phi_3 \in \mathcal{R}_{\alpha,\beta}$ be the non zero determinant of a submatrix of $\mathcal{H}$ of rank $2nd$ multiplied by an appropriate power of Φ_2.

We identify the set of ordered pairs (p, q) of bivariate polynomials p and q of degree at most d in x and n in y with the space $\mathsf{K}^{2(n+1)(d+1)}$, of which $\Phi = \Phi_1\Phi_2\Phi_3 \in \mathcal{R}_{\alpha,\beta}$ defines a hypersurface. Using the special polynomials $\bar{p}$ and $\bar{q}$ of Section 3 we have proven that Φ is not identically zero, and the construction of Φ ensures appropriate properties for computing the resultant outside the hypersurface. Indeed, $\Phi_1 \neq 0$ ensures that $\det Q$ is not a strict divisor or the resultant; $\Phi_2 \neq 0$ provides invertibility of S and strict properness of S^{-1}; $\Phi_3 \neq 0$ leads to denominators Q with small column degrees.

Proposition 4.1. *Let p and q in $\mathsf{K}[x, y]$ be of degree d in x and n in y. If the coefficients of p and q do not form a zero of $\Phi = \Phi_1\Phi_2\Phi_3$ then, for $1 \leq m \leq n$ and X, Y as in (8), the Sylvester matrix S of p and q satisfies*

$$X^T S(x)^{-1} Y = Q_L(x)^{-1} R_L(x) = R(x)Q(x)^{-1} \in \mathsf{K}(x)^{m\times m}$$

for some matrices Q_L, R_L, R and Q of degree at most $\delta = 2\lceil n/m\rceil d$. When the latter descriptions are taken minimal the denominator Q satisfies $\det Q = s\,\mathrm{Res}_y(p, q)$ for some non zero $s \in \mathsf{K}$.

For more insight about the hypersurface to avoid in the space of bivariate polynomials, we note that the (total) degree of Φ in the $\alpha_{i,j}$'s and $\beta_{i,j}$'s is dominated by the degree of Φ_3. The numerators in $\mathcal{H}$ have degree $O(n\delta)$, hence $\deg\Phi_3$ is $O(n\delta \times m\delta)$ and $\deg\Phi$ is $O(n^3d^2/m)$.

The degree bound $2\lceil n/m\rceil d$ in x we consider for Q is sufficient for our purpose since in $O(nd/m)$, and is generically the smallest possible value when m divides n exactly. We note however that a different proof could certainly lead to the sharper bound $\lceil 2nd/m\rceil$.

In the resultant algorithm we will also use the fact that S^{-1} and H are power series since generically $\det S(0) \neq 0$. Proposition 4.1 and Lemma 2.5 then give the following.

Corollary 4.2. *Let p and q generic in $\mathsf{K}[x, y]$ be of degree d in x and n in y. Take $1 \leq m \leq n$, $\delta = 2\lceil n/m\rceil d$, and X, Y as in (8). The first m rows of the m columns of degree at most δ of an order basis of $[X^T S^{-1} Y \ -I_m]$ with order $2\delta + 1$ define a matrix Q such that $\det Q$ is $\mathrm{Res}_y(p, q)$ up to the leading coefficient.*

Note that the specialization $\bar{p}(x, y)$ and $\bar{q}(x, y)$ of Section 3 (used for bounding the generic degree) can be shifted to $\bar{p}(x + \alpha, y)$ and $\bar{q}(x + \alpha, y)$ with an appropriate $\alpha \in \mathsf{K}$ in order to also satisfy also the power series assumption. A different choice of algorithm for the reconstruction could rely simply on the properness of H and handle its expansion at $x = \infty$; see the discussion before Lemma 2.5.

5 TOEPLITZ-LIKE MATRICES

The resultant algorithm given in next section works by reconstructing a matrix fraction from its expansion. In this section we see how to compute a ΣLU (Toeplitz-like) representation [25] of the inverse of a Sylvester matrix. This representation will allow to minimize the cost of the expansion step. For the resultant computation, operations will be on truncated power series with adequate invertibility conditions, hence below we simply consider we are over a field K.

The Sylvester matrix can be seen as a block Toeplitz matrix: after a column permutation, with rectangular 1×2 blocks. As stated in the introduction an appropriate representation of S^{-1} over a field could therefore be computed using the randomized divide and conquer approaches in [27] and [3, Chap. 2, Sec. 13]. For the specific Sylvester matrix case we detail below an alternative (deterministic) solution. Using widely used techniques we derive an explicit formula for S^{-1}; see (16). The formula is in the spirit of those in [31, 37] and slightly more compact (since specific to the Sylvester case). Then we briefly recall how matrix Padé approximation [2] and the half-gcd algorithm can be combined, as suggested in [31], with the formula, for computing a ΣLU representation of S^{-1}. About the relation between the extended Euclidean scheme and structured matrices the reader may also refer to [3, Chap. 2, Sec. 9].

For two coprime univariate polynomials p and q, consider their Sylvester matrix $S \in \mathsf{K}^{(2n)\times(2n)}$. Following the definitions and results of [25], we let Z be the lower shift matrix (ones on the subdiagonal and zeroes elsewhere) and J be the reversal matrix (ones on the antidiagonal and zeroes elsewhere). Since $\operatorname{rank}(S - ZSZ^T) \leq 2$, the (+)-displacement rank of S is $\alpha \leq 2$, therefore we can write

$$S^{-1} = \sum_{i=1}^{2} L_i U_i \in \mathsf{K}^{(2n)\times(2n)}$$

for some lower triangular Toeplitz matrices L_1 and L_2 and some upper triangular Toeplitz matrices U_1 and U_2. Applying the techniques in [21] and [3, Chap 2, Sec. 11] the L_i's and the U_i's can be expressed as Krylov matrices, using iterated powers of Z. For a square matrix M we define the displacement operator

$$\varphi(M) = MZ - ZM. \tag{13}$$

Let e_i denote the ith canonical vector. For the Sylvester matrix one can write $\varphi(S) = \mathcal{V}\mathcal{W}^T$, where $\mathcal{V}$ and $\mathcal{W}$ are matrices with two columns in $\mathsf{K}^{2n\times 2}$ defined by:

$$\varphi(S) = \mathcal{V}\mathcal{W}^T = \begin{bmatrix} Se_{n+1} - Z^n Se_1, -Z^n Se_{n+1} \end{bmatrix} \begin{bmatrix} e_n^T \\ e_{2n}^T \end{bmatrix}.$$

Multiplying $\varphi(S)$ in (13) by S^{-1} on the left and the right we obtain

$$\varphi(S^{-1}) = \mathcal{X}\mathcal{Y}^T, \text{ with } \mathcal{X} = -S^{-1}\mathcal{V}, \text{and } \mathcal{Y}^T = \mathcal{W}^T S^{-1}.$$

Using [3, Chap. 2, Theorem 11.3] and the notation $\mathcal{X} = [x_1, x_2]$ and $\mathcal{Y} = [y_1, y_2]$, it follows that

$$S^{-1} = \mathcal{L}(JS^{-T}e_{2n}) - \sum_{i=1}^{2} \mathcal{U}(ZJx_i)\,\mathcal{L}(Jy_i), \tag{14}$$

where for a vector v of dimension $2n$, $\mathcal{L}(v)$ and $\mathcal{L}(v)$ are the square Krylov matrices of dimension $2n$:

$$\mathcal{L}(v) = [v, Zv, \ldots, Z^{2n-1}v],\ \mathcal{U}(v) = \mathcal{L}(v)^T.$$

Note that since Z is the lower-shift matrix, then $\mathcal{L}(\cdot)$ is Toeplitz lower triangular and $\mathcal{U}(\cdot)$ is Toeplitz upper triangular. Using that $J\mathcal{U}(\cdot)J = \mathcal{V}(\cdot)$, from (14) we derive

$$JS^{-1}J = \mathcal{U}(JS^{-T}e_{2n}) - \sum_{i=1}^{2} \mathcal{L}(ZJx_i)\,\mathcal{U}(Jy_i),$$

and since $y_2 = S^{-T}e_{2n}$,

$$JS^{-1}J = -\mathcal{L}(ZJx_1)\,\mathcal{U}(Jy_1) + \mathcal{L}(-ZJx_2 + I_{2n})\,\mathcal{U}(Jy_2).$$

Rewriting the above for $\varphi(JSJ)$ and $\varphi(JS^{-1}J)$ we obtain:

Proposition 5.1. (Deduced from [21] and [3, Chap. 2, Sec. 11].) *Let p and q be coprime univariate polynomial over* K *of degree n, let $S \in \mathsf{K}[x]^{(2n)\times(2n)}$ be their Sylvester matrix. Define the following vectors in K^{2n}:*

$$s = [-q_{n-1}, \ldots, -q_1, p_n - q_0, p_{n-1}, \ldots, p_0]^T,$$
$$t = [-p_{n-1}, \ldots, -p_1, -p_0, 0, \ldots, 0]^T,$$

and

$$x_1 = S^{-1}s, x_2 = S^{-1}t, y_1 = S^{-T}e_{n+1}, y_2 = S^{-T}e_1. \tag{15}$$

The inverse of the Sylvester matrix then satisfies:

$$S^{-1} = \sum_{i=1}^{2} L_i U_i = \mathcal{L}(Zx_1)\,\mathcal{U}(y_1) + \mathcal{L}(Zx_2 + I_{2n})\,\mathcal{U}(y_2). \tag{16}$$

Representation (16) is slightly more compact than the representations in [31, 37] for general block-Toeplitz matrices. The latter representations could however be compressed into (16) using for instance the solutions of [27, Prop. 4] and [3, Chap. 2, Prob. 2.11b].

We use (16) for representing S^{-1} with $O(n)$ elements of K that are given by the first columns of L_1 and L_2, and first rows of U_1 and U_2. The latter vectors are obtained by solving the linear systems in (15). The first two equations in (15) can be rewritten as polynomial Diophantine equations (for instance see [16, Chap. 4]), and the last two as simultaneous Padé approximation problems [31] that can be solved at essentially the cost $O(\mathsf{M}(n)\log n)$ of the half-gcd [2].

6 THE GENERIC RESULTANT ALGORITHM

The resultant algorithm is given at Figure 1. We use FFT-based polynomial and truncated power series arithmetic [16, Chap. 8,9]. Power series are all truncated at the same order, for a rational matrix M we use $\tilde{M}$ to denote the truncated expansion. For the moment let the blocking factor be $m = \lceil n^\sigma \rceil$ for some $\sigma \geq 0$.

Generically the power series expansions are well defined, in particular $\det S(x) \not\equiv 0 \mod x$; see the comment before Corollary 4.2. Using Proposition 5.1 over truncated power series modulo $x^{2\delta+1}$, Step 3 has cost $(n \times \delta)^{1+o(1)} = (n^{2-\sigma}d)^{1+o(1)}$. We then proceed with the first m entries of the first columns or rows of the truncated L_i's and U_i's. Since the multiplication $m \times m$ Toeplitz matrix times vector is done over a field with $O(\mathsf{M}(m))$ operations, Step 4 has cost $(m^2 \times \delta)^{1+o(1)} = (n^{1+\sigma}d)^{1+o(1)}$. Using the order basis algorithm of [2, 19], Step 5 can be performed in time $(m^\omega \times \delta)^{1+o(1)} = (n^{1+(\omega-1)\sigma}d)^{1+o(1)}$. Using Corollary 4.2 the matrix Q is well defined, its determinant can then be computed using dense $m \times m$ linear algebra for polynomial matrices of degree δ. The cost of Step 6 is bounded by $(m^\omega \times \delta)^{1+o(1)} = (n^{1+(\omega-1)\sigma}d)^{1+o(1)}$ using the randomized algorithm of [48] or the deterministic one in [32]. The correctness of the algorithm follows from Corollary 4.2. By equalizing $2-\sigma$ and $1+(\omega-1)\sigma$ we take $\sigma = 1/\omega$ for minimizing the overall cost, and Theorem 1.1 is proven.

Algorithm *Generic block resultant for p and q in $\mathsf{K}[x, y]$.*

Input: generic polynomials p and q of degree d in x and n in y.
Output: the resultant $\mathrm{Res}_y(p, q)$ of p and q with respect to y.

/ The ˜ notation stands for truncated power series matrices */*

1. $m := \lceil n^{1/\omega} \rceil$. */* Blocking factor */*
2. $\delta = 2\lceil n/m \rceil d$. */* Intermediary degree bound */*
3. */* Expansion step 1. Structured representation of the inverse */*
   Compute the first columns of $\tilde{L}_1$ and $\tilde{L}_2$, and first rows of $\tilde{U}_1$ and $\tilde{U}_2$ using Proposition 5.1 such that
   $$S(x)^{-1} \equiv \sum_{i=1}^{2} \tilde{L}_i(x)\tilde{U}_i(x) \mod x^{2\delta+1} \in \mathsf{K}[x]^{(2n)\times(2n)}.$$
4. */* Expansion step 2. Submatrix of S^{-1} */*
   Using the $m \times m$ Toeplitz submatrices $\tilde{L}_i^{(m)}$'s and $\tilde{T}_i^{(m)}$'s of the $\tilde{L}_i$'s and $\tilde{U}_i$'s such that
   $$\tilde{H}(x) \equiv \sum_{i=1}^{2} \tilde{L}_i^{(m)}(x)\tilde{T}_i^{(m)}(x) \mod x^{2\delta+1},$$
   compute
   $$\tilde{H}(x) \equiv H(x) \equiv X^T S(x)^{-1} Y \mod x^{2\delta+1} \in \mathsf{K}[x]^{m\times m}.$$
5. */* Matrix fraction reconstruction, see Corollary 4.2 */*
   Compute an order basis $P(x) \in \mathsf{K}[x]^{2m\times 2m}$ of $[\tilde{H}(x) \ -I_m]$ with order $2\delta + 1$.
   $Q(x) :=$ the first m rows of the m columns of $P(x)$ of degree $\leq \delta$.
6. */* Determinant computation */*
   $S_d :=$leading matrix of S, $S_d \in \mathsf{K}^{(2n)\times(2n)}$.
   $s := \det S_d \in \mathsf{K} \setminus \{0\}$.
   $r(x) := \det Q(x)$.
   Return $(s/r_{2n}) \cdot r(x)$.

Fig. 1. Computation of the resultant.

Since we have genericity assumptions we could rely either on randomized or deterministic algorithms at every stages of the resultant algorithm. For example, thanks to the genericity, the randomized approaches of [27] and [3, Chap. 2, Sec. 13] would behave deterministically at Step 3. The same is true for the randomized algorithm of [48] for the final determinant computation. We remark however that our complexity bound has been derived using deterministic solutions. As far as we know, the use of the best known randomized strategies would lead to an improvement restricted to log factors.

7 EXTENSIONS

We present here some consequences of our approach.

Generic bivariate ideals. By Proposition 4.1 we have $X^T S^{-1} Y = RQ^{-1}$ with $\det Q = s \det S$ where $s \in \mathsf{K} \setminus \{0\}$. If Q' is the right denominator of a minimal description of $S^{-1}Y$ then it is also a denominator for $X^T S^{-1} Y$ and by Lemma 2.2 we know that Q' is a multiple of Q. Applying Lemma 2.1 to left and right descriptions of $S^{-1}Y$ we see that $\det Q'$ divides

$\det S$, hence $\det Q$, it follows that $Q' = QU$ with U unimodular, and Q is also a denominator for a minimal description of $S^{-1}Y$. Consequently, for some some polynomial matrix P one has $S(x)P(x) = \left[0, \ldots, 0, Q(x)^T\right]^T \in \mathsf{K}[x]^{(2n)\times m}$. Thanks to the form of S, the polynomials $\sum_{i=1}^{m} q_{i,j}(x)y^{m-i}$, $1 \le j \le m$, constructed from the entries of Q, are in the ideal $\mathcal{I}$ generated by p and q. If by right unimodular equivalence we compute the Hermite normal form G of Q as a lower triangular matrix, the last two columns of G give two polynomials $\gamma_1(x, y) = g_{m-1,m-1}(x)y + g_{m,m-1}(x)$ and $\gamma_2(x) = g_{m,m}(x)$ in the ideal $\mathcal{I}$. The polynomial γ_2 divides $\det Q$, hence $\det S$. Further, from the Hermite form of the Sylvester matrix of $\bar{p}$ and $\bar{q}$ in Section 3 with $m = 1$, one can see that generically one must have $\gamma_2 = s \det S$ for $s \neq 0 \in \mathsf{K}$, which implies that $g_{m-1,m-1}$ is an element in K.

Since the Hermite form is normalized with monic diagonal entries we have $\gamma_1(x, y) = y - g_{m,m-1}(x)$ and $\gamma_2(x) = s \det S$. This is also the fact that generically $\mathcal{I}$ is generated by $\det S = \mathrm{Res}_y(p, q)$ and a polynomial $y - g(x)$ that form a Gröbner basis of $\mathcal{I}$ for the lexicographic order with $y > x$ [18].

We see that the computation of the Hermite form of Q provides a Gröbner basis of the ideal generated by p and q. Using the notations of previous sections, Q is $m \times m$ of degree δ, its Hermite form can be computed in $(m^\omega \times \delta)^{1+o(1)}$ operations using the randomized algorithm of [20] or the deterministic one in [32]. The complexity bound of Theorem 1.1 remains valid for the computation of a Gröbner basis of the ideal $\langle p, q\rangle$.

Characteristic polynomial of a generic structured matrix. Parallel to the construction of Section 3, for $1 \le m \le n$ let us consider the two matrices: $B' \in \mathsf{K}[x]^{m\times n}$ such that $B'_{i,i} = x$ and $B'_{i,i+n-m} = -1$ for $1 \le i \le m$, and $B'_{i,j} = 0$ otherwise; $B'' \in \mathsf{K}[x]^{(n-m)\times n}$ such that $B''_{i,i} = 1$ and $B''_{i,i+m} = x$ for $1 \le i \le n - m$, and $B''_{i,j} = 0$ otherwise. Consider also $\bar{T}(x) = \left[(B'(x))^T,\ (B''(x))^T\right]^T \in \mathsf{K}[x]^{n\times n}$. One can check that $X^T\bar{T}^{-1}Y$ can be described in a manner similar to (9). The fraction descriptions now have degree bounded by $\delta = \lceil n/m\rceil$. Since $\bar{T} = xI_n - \bar{T}_0$, the determinant of $\bar{T}$ is the characteristic polynomial of $\bar{T}_0$, and note that $\bar{T}_0$ is Toeplitz.

This shows that for a generic Toeplitz matrix T_0, a modification of the algorithm of Section 6 using a ΣLU representation of $T^{-1} = (xI_n - T_0)^{-1}$ will compute the characteristic polynomial within the complexity bound of Theorem 1.1 with $d = 1$. One may expect to have analogous results for more general classes of structured matrices [6, 14].

Characteristic polynomials in univariate quotient algebras. We consider a slight modification of $\bar{p}$ and $\bar{q}$ of Section 3. For $1 \le m \le n$ take $\bar{p}(x, y) = y^n + y^m$ (or $\bar{p}(x, y) = y^n$ for $m = n$ in characteristic 2) and $\bar{q}(x, y) = y^n + x$. With $\bar{S}$ the Sylvester matrix of $\bar{p}$ and $\bar{q}$ the fraction $X^T\bar{S}^{-1}Y$ can be described similarly to (9), with $\delta = \lceil n/m\rceil$.

Note that we are not in a generic Sylvester case since for instance the coefficient matrix of degree one of $\bar{S}$ is singular. We are however in a generic situation for the special resultant of polynomials $x - a(y)$ and $g(y)$ of degree n in y. The arguments of Sections 3 and 4 can be extended to the situation here. The special Sylvester matrix has dimensions $(2n) \times (2n)$ and for generic a and g the resultant has degree n. Modified Φ_2 and Φ_3 have to be used for an appropriate extension of Proposition 4.1, H remains strictly proper (which is not true anymore for S^{-1} in general).

This shows that the complexity bound $O(n^{1.58})$ of Theorem 1.1 with $d = 1$ is valid generically for computing the characteristic polynomial χ of the multiplication by a in $\mathcal{A} = \mathsf{K}[y]/\langle g(y)\rangle$.

Previously existing general resultant algorithms compute the characteristic polynomial in $\mathcal{A}$ at cost $n^{2+o(1)}$. However, in this special resultant case, the minimal and characteristic polynomial problems, respectively in [46] and [5], are reduced to the modular power projection problem, and by duality to the modular composition problem [28, 46]. With an adaptation of the composition algorithm of [7] this leads to algorithms using $n^{1.5+o(1)} + O(n^{(\omega+1)/2})$ operations in K for the minimal and characteristic polynomials [5, 46]. Note that for small characteristic fields the approach in [5] asks

the root multiplicity of χ to be less than the characteristic of K; small fields can be handled using p-adic techniques such as in [33].

Our bound $n^{2-1/\omega+o(1)}$ shows that in the generic case with $\omega > 2$ the composition algorithm of [7] can be bypassed. As evoked above, composition is dual to power projection. The algorithms of [5, 46] rely on particular K-linear maps $\pi : \mathcal{A} \to \mathsf{K}$ for the projections. The same is true for our algorithm. Since special maps can be used, the characteristic polynomial problem may not be as difficult as the general composition problem.

Acknowledgements. The author is grateful to the referees for their careful readings and helpful comments; to E. L. Kaltofen for his motivating questions; to J.-C. Faugère, A. Galligo, V. Neiger, M. Safey El Din and B. Salvy for rich discussions on Gröbner bases.

REFERENCES

[1] J. Abbott, M. Bronstein, and T. Mulders. 1999. Fast deterministic computation of determinants of dense matrices. In *Proc. ISSAC*. ACM Press.
[2] B. Beckermann and G. Labahn. 1994. A uniform approach for the fast computation of matrix-type Padé approximants. *SIAM J. Math. Anal. Appl.* 15, 3 (1994).
[3] D. Bini and V.Y. Pan. 1994. *Polynomial and matrix computations*. Birkhäuser.
[4] R.R. Bitmead and B.D.O. Anderson. 1980. Asymptotically fast solution of Toeplitz and related systems of linear equations. *Linear Algebra Appl.* 34 (1980).
[5] A. Bostan, P. Flajolet, B. Salvy, and É. Schost. 2006. Fast computation of special resultants. *J. Symbolic Computation* 41, 1 (2006).
[6] A. Bostan, C.-P. Jeannerod, C. Mouilleron, and É. Schost. 2017. On matrices with displacement structure: generalized operators and faster algorithms. *SIAM J. Math. Anal. Appl.* 38, 3 (2017).
[7] R.P. Brent and H.T. Kung. 1978. Fast algorithms for manipulating formal power series. *J. ACM* 25, 4 (1978).
[8] W.S. Brown. 1971. On Euclids algorithm and the computation of polynomial greatest common divisors. *J. ACM* 18, 4 (1971).
[9] D.G. Cantor and E. Kaltofen. 1991. On fast multiplication of polynomials over arbitrary algebras. *Acta Informatica* 28, 7 (1991).
[10] G.E. Collins. 1971. The Calculation of Multivariate Polynomial Resultants. *J. ACM* 18, 4 (1971).
[11] D. Coppersmith. 1997. Rectangular matrix multiplication revisited. *J. Complexity* 13 (1997).
[12] D. Coppersmith and S. Winograd. 1990. Matrix multiplication via arithmetic progressions. *J. Symbolic Computation* 9, 3 (1990).
[13] D. Cox, J. Little, and D. O'Shea. 1998. *Using Algebraic Geometry*. Springer-Verlag, New-York. 2nd edition 2005.
[14] C. De Sa, A. Cu, R. Puttagunta, C. Ré, and A. Rudra. 2018. A two-pronged progress in structured dense matrix vector multiplication. In *Proc. ACM-SIAM SODA*.
[15] W. Eberly, M. Giesbrecht, and G. Villard. 2000. Computing the determinant and Smith form of an integer matrix. In *Proc. FOCS*. IEEE.
[16] J. von zur Gathen and J. Gerhard. 1999. *Modern Computer Algebra*. Cambridge University Press. Third edition 2013.
[17] J. von zur Gathen and T. Lücking. 2003. Subresultants revisited. *Theoretical Computer Science* 297, 1-3 (2003).
[18] P. Gianni and T. Mora. 1987. Algebraic solution of systems of polynomial equations using Gröbner bases. In *Proc. AAECC (LNCS 536)*.
[19] P. Giorgi, C.-P. Jeannerod, and G. Villard. 2003. On the complexity of polynomial matrix computations. In *Proc. ISSAC*. ACM Press.
[20] S. Gupta. 2011. *Hermite forms of polynomial matrices*. Master Thesis. University of Waterloo. http://hdl.handle.net/10012/6108
[21] G. Heinig and K. Rost. 1984. *Algebraic Methods for Toeplitz-like Matrices and Operator*. Springer, Birkhäuser Basel.
[22] X. Huang and V.Y. Pan. 1998. Fast rectangular matrix multiplication and applications. *J. of Complexity* 14, 2 (1998).
[23] C.-P. Jeannerod, V. Neiger, É. Schost, and G. Villard. 2016. Fast computation of minimal interpolation bases in Popov form for arbitrary shifts. In *Proc. ISSAC*. ACM Press.
[24] T. Kailath. 1980. *Linear Systems*. Prentice-Hall.
[25] T. Kailath, S.Y. Kung, and M. Morf. 1979. Displacement ranks of matrices and linear equations. *J. Mathematical Analysis and Applications* 68, 2 (1979).
[26] E. Kaltofen. 1992. On computing determinants without divisions. In *Proc. ISSAC*. ACM Press.
[27] E. Kaltofen. 1994. Asymptotically fast solution of Toeplitz-like singular linear systems. In *Proc. ISSAC*. ACM Press.
[28] E. Kaltofen. 2000. Challenges of symbolic computation: my favorite open problems. *J. Symbolic Computation* 29, 6 (2000).
[29] E. Kaltofen and G. Villard. 2005. On the complexity of computing determinants. *Computational Complexity* 13, 3 (2005).
[30] D. E. Knuth. 1971. The analysis of algorithms. In *Actes, Congrès int. math. (Nice, 1970)*. Tome 3. Gauthier-Villars, 269–274. http://perso.ens-lyon.fr/gilles.villard/BIBLIOGRAPHIE/Knuth-ICM-1970.pdf
[31] G. Labahn. 1992. Inversion components of block Hankel-like matrices. *Linear Algebra Appl.* 177 (1992).
[32] G. Labahn, V. Neiger, and W. Zhou. 2017. Fast deterministic computation of the Hermite normal form and determinant of a polynomial matrix. *J. Complexity* 42 (2017).
[33] P. Lairez and T. Vaccon. 2016. On p-adic differential equations with separation of variables. In *Proc. ISSAC*. ACM Press.

[34] F. Le Gall. 2014. Powers of Tensors and Fast Matrix Multiplication. In *Proc. ISSAC*. ACM Press.

[35] F. Le Gall and F. Urrutia. 2018. Improved rectangular matrix multiplication using powers of the Coppersmith-Winograd tensor. In *Proc. ACM-SIAM SODA*.

[36] G. Lecerf. 2017. *On the complexity of the Lickteig-Roy subresultant algorithm.* HAL report. CNRS & École Polytechnique.

[37] L. Lerer and M. Tismenetsky. 1986. Generalized Bezoutian and the inversion problem for block matrices I. General scheme approximants. *Integral Equations and Operator Theory* 9, 6 (1986).

[38] T. Lickteig and M.-F. Roy. 1996. Cauchy index computation. *Calcolo* 33,3-4 (1996).

[39] R. T. Moenck. 1973. Fast computation of GCDs. In *Proc. STOC*. ACM Press.

[40] G. Moroz and É. Schost. 2016. A fast algorithm for computing the truncated resultant. In *Proc. ISSAC*. ACM Press.

[41] V. Pan. 1988. Computing the determinant and the characteristic polynomial of a matrix via solving linear systems of equations. *Inf. Process. Lett.* 28, 2 (1988).

[42] V. Pan. 1992. Parametrization of Newtons iteration for computations with structured matrices and applications. *Comp. Math. Appl.* 24, 3 (1992).

[43] D. Reischert. 1997. Asymptotically fast computation of subresultants. In *Proc. ISSAC*. ACM Press.

[44] J. Rifà and J. Borrell. 1991. Improving the time complexity of the computation of irreducible and primitive polynomials in finite fields. In *Proc. AAECC(LNCS 539).*

[45] A. Schönhage. 1971. Schnelle Berechnung von Kettenbruchentwicklungen. *Acta Informatica* 1 (1971).

[46] V. Shoup. 1994. Fast construction of irreducible polynomials over finite fields. *J. Symbolic Computation* 17, 5 (1994).

[47] V. Shoup. 1995. A New Polynomial Factorization Algorithm and its Implementation. *J. Symbolic Computation* 20, 4 (1995).

[48] A. Storjohann. 2003. High-order lifting and integrality certification. *J. Symbolic Computation* 36, 3-4 (2003).

[49] A. Storjohann. 2005. The shifted number system for fast linear algebra on integer matrices. *J. Complexity* 21, 4 (2005).

[50] J.A. Thiong Ly. 1989. Note for computing the minimun polynomial of elements in large finite fields. In *Proc. Coding Theo. App. (LNCS 388).*

[51] G. Villard. 1997. Further analysis of Coppersmiths block Wiedemann algorithm for the solution of sparse linear systems. In *Proc. ISSAC*. ACM Press.

[52] G. Villard. 1997. *A study of Coppersmith's block Wiedemann algorithm using matrix polynomials.* RR 975 IM IMAG Grenoble.

[53] D. Wiedemann. 1986. Solving sparse linear equations over finite fields. *IEEE Trans. Information Theory* 32, 1 (1986).

[54] W. Zhou and G. Labahn. 2012. Efficient Algorithms for Order Basis Computation. *J. Symbolic Computation* 47, 7 (2012).